Official URL: https://doi.org/10.1016/j.ifacol.2018.06.193

Tahiri, Ayoub and Ladevèze, David and Chiron, Pascale and Archimède, Bernard Improving the characterization of initial conditions for streamflow prediction using a precipitation reconstruction algorithm. (2018) In: 1st IFAC Workshop on Integrated Assessment Modelling for Environmental Systems, (IAMES 2018), 10 May 2018 - 11 May 2018 (Brescia, Italy).

Improving the characterization of initial conditions for streamflow prediction using a precipitation reconstruction algorithm

Ayoub TAHIRI *. David LADEVEZE*. Pascale CHIRON. Bernard ARCHIMEDE****

**Compagnie d'Aménagement des Coteaux de Gascogne, Tarbes, France*
(e-mail: a.tahiri@cacg.fr ; d.ladeveze@cacg.fr).
***Ecole Nationale d'Ingénieurs de Tarbes, Tarbes, France*
(e-mail: pascale.chiron@enit.fr ;bernard.archimede@enit.fr)

Abstract: Hydrologic forecasts derive their skill from knowledge of initial conditions at the forecast date, climate forecast, model structure and parameters. Uncertainty on the initial conditions has as much influence as uncertainty on the weather forecasts on the hydrologic forecasts for some watersheds. Initial conditions depend on several parameters: evapotranspiration, soil composition and mainly former rain events which are measured by rain gauges or radars. Precipitation measures often show uncertainties or even data gaps, and thus, the evolution of the soil states is unknown. The initial conditions can only be determined by following up the evolution of the variable states. The measured discharge runoff is the only available reliable data and thus, that information can be used to determine the variable states, by the inversion of the rainfall-runoff model. This study proposes a post-processing method that adjust the initial conditions using measured discharge runoff at the outlet of a watershed. The heuristic is applied on the Echez watershed, and the effectiveness of the method is illustrated thanks to a comparison of the results obtained with the measured observation during an analysis period falling out after the forecast date.

Keywords: Hydrology backward, precipitation reconstruction, hydrologic forecast, initial conditions.

1. INTRODUCTION

Water supply outlooks, or volume of runoff, are communally used by water managers in order to predict water supplies, determining industrial and agricultural water allocations, and operating reservoirs for multiple uses such as hydropower and flood control. Therefore, each water manager should take advantage of a tool that can predict the evolution of the water resource in short and long term in order to adapt his management strategy.

A hydrologic model is a mathematical model describing the rainfall-runoff process at the scale of a catchment area, drainage basin or a watershed. The inputs of the model vary from a model to another (precipitation, evapotranspiration, soil permeability, ...) and the output is the discharge runoff at the watershed outlet. The differences between the simulated flow rates and the observed flow rates represent the errors of the model.

Every hydrologic model has a number of parameters that need to be calibrated based on the available observations, so that the model can simulate the catchment hydrological behaviour as closely as possible. The model calibration process consists in varying the model parameters until the measured flow corresponds to the model outflow using a large former data set. Different efficiency criteria for hydrological model calibration are mentioned in [Nash and Sutcliffe, 1970] and [Krause et al, 2005]. The type and the number of parameters are different from a model to another. Some parameters may be related to physical characteristics of the watershed, others are abstract quantities (storage capacity of a soil tank, etc.).
Hydrologic models can be classified into different groups depending on their conception and the nature of the expressions defining the relationships between the inputs and outputs: deterministic or stochastic; discrete in time or continuous; physical, empirical or conceptual; discrete in space or global…

A conceptual model allows to represent the main processes of the rain-flow relationship without describing the physical laws governing the processes involved. This type of model generally consists in interconnected reservoirs, in which the level increases and decreases over time and represents the different hydrological compartments of the watersheds.

In the current study, we focus on conceptual, deterministic and global models, and we assume that the models are calibrated. The advantage of a conceptual approach is that the model is much simpler from a mathematical point of view. Thus, hydrologic processes are estimated with simple equations rather than solving governing partial differential equations and so, setting and calibration is easier [Aghakouchak and Habib, 2010].

In order to execute a simulation on a hydrologic model, it is necessary to specify the initial conditions of the simulation, which represent the variables states at the first-time step of the simulation. The initial hydrologic conditions have a strong impact on the prediction of cumulative runoff and soil moisture [Shukla and Lettenmaier, 2011]. Errors on the initial states have as much influence on the quality of the flow prediction than those related to weather forecast [Kirchner, 2009]. In fact, in the process transforming rainfall to runoff at the catchment scale, former rains events influent strongly the response of this basin via the saturation of the soil. When the soil is saturated, the watershed tends to respond rapidly and intensively to rainfall, while when it is dry, the watershed absorbs most of the rainfall. The knowledge of the antecedent moisture degree has been a major challenge for hydrological prediction, mainly for two reasons: (1) it's difficult to exactly estimate catchment soil moisture status through time, by either measurement or modelling, and (2) it's difficult to determine the functional relationship between this antecedent soil moisture and the runoff which induced it. Therefore, improvement in knowledge of the initial hydrologic conditions would improve the streamflow prediction [Shukla and Lettenmaier, 2011].

Initial conditions depend on several parameters: evapotranspiration, soil composition and mainly former rain events which are measured by rain gauges or radars. Many reasons can lead to uncertainties when estimating rainfall such as rain gauge technical ability to measure severe rainfall intensities, poor spatial or temporal resolution rain gauges sampling [Sarann et al, 2012]. Precipitation measurements often show uncertainties or even data gaps, and thus, the evolution of the states is broken. The initial conditions can only be determined by following up the evolution of the variables states. The measured discharge runoff is the only available reliable data and thus, that information can be used to determine the variable states, by inverting the rainfall-runoff model.

Recent contributions to the literature have raised the question to figure out the uncertainty of measured rainfall. Some studies suggested methods to better take into account rainfall uncertainties during the calibration of rainfall-runoff models. These studies suggested to consider rainfall at each time step as a model parameter that need to be calibrated as usual model parameters. However, the number of parameters to be calibrated is large, thus, other studies proposed the use of correction factors for rainfall series in order to reduce the number of parameters to be calibrated. [Kuczera et al. 2006] [Vrugt et al. 2008] solve the input uncertainty in hydrologic modelling, using a markov chain monte carlo simulation.

[Kirchner 2009] represented the catchment as a simple first-order nonlinear dynamical system which is thus, invertible; one needs only measured streamflow fluctuations as input to calculate (P-E), with P: the precipitations and E: the evapotranspiration. A similar approach was used by [Teuling et al. , 2010], [Krier et al. , 2012] and [Herrnegger et al. 2014]. [Michon 2015] extended Krichner's work to non-analytical rainfall-runoff models. The suggested heuristic inversion method is based on a Quasi-Newton algorithm and showed to be able to identify both hourly rainfall time series and rainfall-runoff model parameters values. The problem with this approach is that the rainfall time series and the parameters are unknown, and so, the quality of the inverted rainfall depends on the capacity of the algorithm on finding precisely the model parameters, since the model is sensitive to the latter.

In this work, we propose a precipitation reconstruction algorithm from the chronicles of measured flow rates in order to estimate the initial states. It consists in finding the hourly rainfall time series which generated the observed flows. The reconstituted precipitations will be reinjected in the hydrologic model and initial conditions will be found. The algorithm developed herein relies on a numerical model coupled with a hydrologic model, and allows the user to generate better hydrologic forecast data.

2. Methodology and mathematical formulation

2.1 Methodology

Reconstituting precipitations consists in determining rain events such that when these events are input in the hydrologic model, the simulated flow rate results approach the observed flow rate. In other words, it consists in finding precipitation values at each time step that minimize the difference between the output of the hydrologic model and the measured flow (see Figure 1).

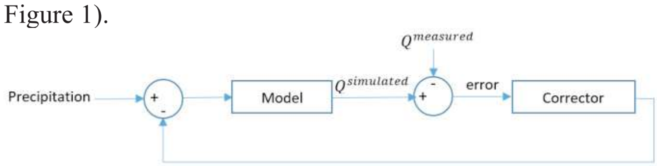

Figure 1: Precipitations identification methodology

The methodology consists mainly in three steps:

1. Generate a precipitation vector.
2. Introduce the precipitation vector as an input for simulation on a calibrated model, and evaluate the error defined as the gap between the simulated flow and the measured one.
3. Integrate this error, and generate a new precipitation vector that will reduce the error.

The complexity of this problem is mainly due to two reasons:

- The model is a black box: The mathematical transformation function and its derivatives are not known and their evaluation is computing time expensive.
- The model has a state representation and thus, a bad estimation of a rain value at instant t, will change the state all over the simulation horizon which make it difficult to find the correct future values.

Let P, $Q^{sim}(P)$ and Q^{meas} denote the precipitation vector, the simulated flow corresponding to vector P and the measured flow respectively. Let $E(P)$ denote the error vector of the model corresponding to the precipitation vector P:

$$E(P) = Q^{meas} - Q^{sim}(P) \tag{1}$$

2.2 Mathematical formulation

An objective function must be defined in order to quantify the difference between the observed flow and the flow simulated by the hydrological model. Several objective functions can be implemented for the reconstruction computation: absolute-value norm, Euclidian norm…

In order to avoid the compensation of opposite signs errors, the $l_1$norm which is given by the sum of the absolute values of the vector elements E(P) is used as an objective function.
The problem of precipitation reconstruction in a hydrologic model can be modelled as a constrained optimization problem, by adding bounds on elements of the precipitation vector.

$$Min_{P \in R^n} \ |E(P)| \tag{2}$$
$$P_t \geq 0 \quad \forall t \in \{1,2, \dots n\} \tag{3}$$

With n represents the time horizon, and P_t the precipitation at instant t.

Resolving the problem of precipitation reconstruction consists in identifying the set of rainfall variables for which the model outputs are the measured flows. This is done by a numerical optimization procedure whose purpose is to determine the values of the input variables producing the vector P while minimizing the objective function $|E(P)|$.

Theoretically, a vector P for which the objective function is zero exists, nevertheless the set of measured flow contains uncertainties. Therefore, the objective function cannot achieve its theoretical minimum value, and the obtained precipitation variables input set does not strictly correspond to the actual one.

2.3 Optimization-Simulation mechanism

Optimization problems in hydrological models are generally complex because of the non-linearity of the mathematical formalizations and because of the models structures. Optimum research strategies can be divided into two categories: local methods and global methods.

Local methods explore in a progressive and evolutionary way the state space from an initial input value. The exploration of space is done in the direction where it is possible to improve the value of the objective function. The procedure can be terminated when it's no longer possible to generate a significant improvement, thereby the set of variables found, corresponds to an optimum of the objective function. These methods can be based on the objective function only or on its derivatives as well. The advantage of these methods is that at each iteration, the objective function gets improved. However, the main drawback of these kind of methods is that they can lead to a local optimum instead of finding the global one.

Global methods explore a large research space and therefore allow to identify, in principle, the optimal set of input variables and avoid local convergence. Among these methods, we can mention: simulated annealing [Vicente, et al, 2003], genetic algorithms [Mitchell, 1998], neural network [Schalkoff, 1997]. These methods use the vectors tested during former iterations to generate the new one. Even if the generated vector doesn't improve the objective function, it is kept in a database and could be used in a future iteration. These methods are time consuming for the first iterations, but speed up as the database get bigger.

The optimization algorithm implemented herein is global, iterative and improves the objective function at each iteration. The end of an iteration is marked with a model simulation. Let P^k and P_t^k denote for an iteration k, the precipitation vector and the precipitation at instant t. The algorithm starts with a null vector $P^0(0,0, \dots ,0)$ and evolves to the optimal vector P^{opt}. At each iteration, the recurrence formula (4) is applied:

$$P^{k+1} = P^k + \beta^k I(P)^k \tag{4}$$

With $I(P)^k$ an indicator defined later, and β^k a correction vector at iteration k.

Every watershed and so, every hydrologic model has its own reaction to a rain event. To determine the response of the hydrologic model to a rain event we consider a database which contains the results of every simulation made during the previous iterations of the algorithm. The algorithm uses the database simulations values to infer the corrections to be made on the precipitation vector at each instant t, so that the next input precipitation vector can improve the output flow. A first simulation with the vector $P'(10\text{mm}, 0, \dots 0)$ is made to initialize the database.

Every single rain event P_t has an effect on the flow all along the concentration time m of the watershed. In other words P_t has an effect on $\{Q_t^{sim}, Q_{t+1}^{sim} \dots Q_{t+m}^{sim}\}$. In order to consider the error over this period, an indicator consisting of the convolution of $f = Q^{sim}(P) - Q^{meas}$ and $g = Q^{sim}(P) - Q^{sim}(P')$ is computed:

$$I(P)[i] = (f * g)[i] = \sum_{k=0}^{k=m} f[i+k]g[k] \tag{5}$$

Vectors f and g represent the error of the simulation and the signature of the model respectively. The indicator (5) is a similarity measure of the two series (f, g) and is defined as a function of the displacement of one relative to the other. In other words, it corresponds to a cross-correlation measure.
The algorithm selects the points that require corrections, i.e. the instants for which the value of $I(P)$ is significant. Let t denote one of those instants, the algorithm looks for rain events in the database which are similar to the one at instant t. Let T denote the set of rain events similar to one at instant t.The similarity measure is based on three standards: the temporal distance between the two events, the cumulative rainfall value, and the quantity of instantaneous rain. Based on these three standards values, a corrector coefficient β_t^k is computed using the empirical formula (6).

$$\beta_t^k = Mean_{i\in T}\left(\frac{(P^k - P^{k-1})[i]}{(I^k - I^{k-1})[i]}\right) \quad (6)$$

The reconstructing algorithm steps are illustrated in the Figure 2.

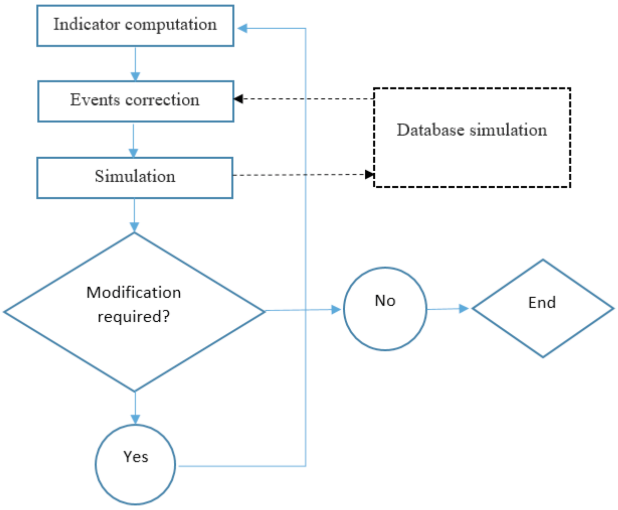

Figure 2: Reconstruction algorithm

At each iteration, the algorithm generates a net rainfall vector which is simulated on a calibrated hydrological model to obtain the flow rates associated. The indicator is then computed in order to estimate the rainfall events values at each time step. The errors, revealed out by a significant indicator values are corrected by modifying the rain event at the corresponding time steps. The correction quantities are computed using the former simulations in the database and the formulas (4), (5) and (6). For the foremost iterations, few simulations exist in the database, hence the corrections proposed are not accurate, but as the database get bigger, the algorithm becomes more efficient and faster.

3. Case study

The algorithm is applied on the GR3H hydrologic model [Fourmigué et al, 2005; Perrin et al, 2007]. The hydrologic model GR3H is a conceptual model. The input is the hourly rain vector on the watershed and the output is the hourly flow at the outlet. The reconstruction algorithm was applied on the Echez watershed. The measuring station used in this study is located in Tarbes, in the south west of France and covers a surface of 233 Km^2.

Firstly, in order to validate the algorithm, we test its ability of finding precipitations that generated the observed flows. In other words, we test the algorithm ability for inverting the model. The validation stage is important, because it's the only way to confirm the robustness of the algorithm.

The validation stage is performed over the period (03/03/2017 to 15/04/2017). Figure 3 represent the hourly reconstituted precipitations, the measured precipitations, the hourly measured flow and the hourly simulated flow using the reconstituted precipitations. The comparison between the two precipitations vector shows a temporal disparity. On the other hand, the total volume of rain measured and reconstituted over the study period is 183.56 mm and 179.4 mm respectively, which corresponds to an error of 2.32%. The origin of this error may be due to different reasons: errors of measurements (precipitations and inflows), the calibration of the GR3H model, or due to the fact that we didn't take into account the evapotranspiration, since the reconstituted precipitation corresponds to the net rainfall.

While comparing the hourly measured flow and the hourly simulated flow it can be seen that the algorithm succeeded to reproduce the measured flows. The error between the simulated flow and the measured one over the validation period (42 days) is 1.5%.

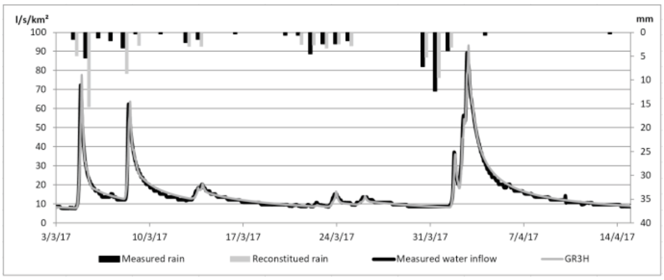

Figure 3: reconstituted and measured data

As a conclusion, the algorithm achieved the reconstruction of the measured flow. For a period of 42 days, which corresponds to a 1008 hourly period, the algorithm had to make only 9 simulations to reach optimality.

After validating the algorithm, we evaluate the contribution of the reconstructing algorithm for the hydrological forecast. Four sequences of forecast at different dates are considered in order to observe the evolution in time of the forecast after reassessing the initial conditions using the precipitation reconstruction algorithm. The different simulations are made with the same precipitation forecast which is obtained retrospectively from the real measured precipitation. For the four sequences, the algorithm uses the observed flow measured prior the starting date of the sequences.

The study covers a 20 days' period (from 26/12/2017 to 15/01/2018). The GR3H model is calibrated, and uses the precipitation forecasts to predict the discharge. For the four sequences, the starting dates of the forecast and the durations are listed in table1.

Table 1: Sequences of forecast

Sequence	A	B	C	D
Starting date	27/01/2017	29/01/2017	01/01/2018	05/01/2018
Period of forecast	2 days	3 days	4 days	10 days

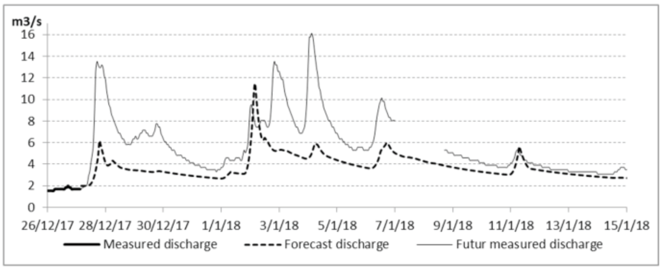

Figure 4: Sequence A

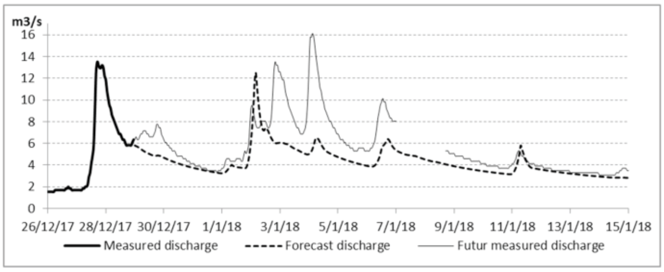

Figure 5: Sequence B

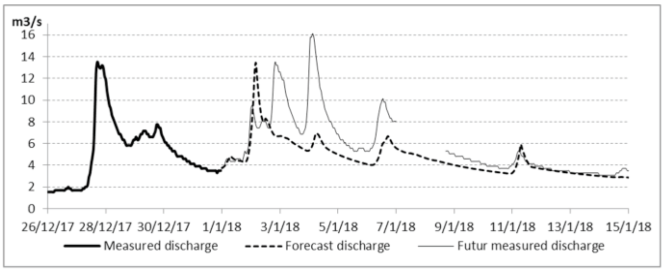

Figure 6: Sequence C

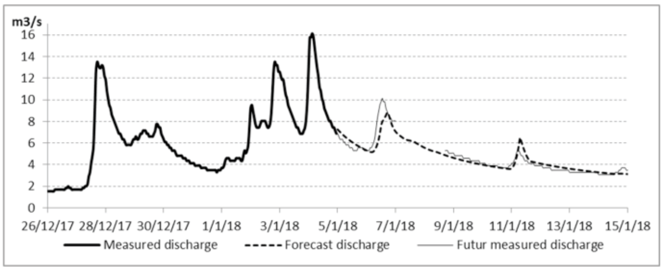

Figure 7: Sequence D

Figures 4, 5, 4, 7 show the evolution in time of the hydrologic forecast. As it is shown, the continuously updating of the initial conditions significantly improved the forecast. We remind that the simulations are made using the same precipitation forecast. Hence, the improvements are only due to the updating of the initial conditions. The errors of forecast are of $35\%, 24\%, 21\%$, and 1% for the sequences A, B, C and D respectively.

For the sequences A, B and C, it's obvious that it's not the best hydrologic forecast, but this is due to the non-representative precipitation measurement.

Figure 7 shows that with a good weather forecast and a calibrated initial conditions, the model can provide good results if the measure station is temporarily failing as is the case between 07/01/2018 and 09/01/2018.

6. CONCLUSIONS

The aim of this paper is to evaluate the initial conditions of a hydrologic model using inflow measurement. An algorithm is proposed to inverse a hydrologic model, and hence determine the initial conditions of the simulation, which represent the variables states at the first-time step of the simulation. The developed heuristic is global, iterative and approach the correct initial conditions at each iteration. On the case study, the proposed method was shown to estimate reasonably well precipitation values that enable the model to reproduce the measured inflow. The capacity of the model to reproduce the measured inflow assure the good follow-up of the variable states and so improve the hydrological forecast. Furthermore, it was shown through the four sequences that the hydrological forecast gets improved after using the inflow measurement to update the initial conditions. One of the perspectives of this work would be the application of the algorithm on different hydrological models in order to draw general conclusions.

REFERENCES

Aghakouchak, A., & Habib, E. (2010). Application of a conceptual hydrologic model in teaching hydrologic processes. International Journal of Engineering Education, 26(4 (S1)).

P. Fourmigué & J. Lavabre. (2005). Prévision de crues avec le modèle conceptuel pluie-débit GR3H. Adaptabilité aux incertitudes sur la pluie. Revue des sciences de l'eau/Journal of Water Science, 18(1), 87-102.

Herrnegger, M., H. P. Nachtnebel, and K. Schulz (2014), From runoff to rainfall: inverse rainfall runoff modelling in a high temporal resolution, Hydrology and Earth System Sciences Discussions,

Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. Water Resources Research, 45(2).

Krause, P., Boyle, D. P., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. Advances in geosciences, 5, 89-97.

Krier, R., P. Matgen, K. Goergen, L. Pfister, L. Hoffmann, J. W. Kirchner, S. Uhlenbrook, and H. H. G. Savenije (2012), Inferring catchment precipitation by doing hydrology back-ward: a test in 24 small and mesoscale catchments in Luxembourg, Water Resources Research, 48(10), n/a–n/a, doi: 10.1029/2011WR010657.

Kuczera, G., D. Kavetski, S. Franks, and M. Thyer (2006), Towards a bayesian total error analysis of conceptual rainfall-runoff models : characterising model error using storm-dependent parameters, Journal of Hydrology, 331(1-2), 161–177, doi :10.1016/ j.jhydrol.2006.05.010.

Michon, T. (2015). Extension du potentiel de la modélisation hydrologique.: inversions heuristiques de modèles pluie-débit pour l'identification des paramètres simultanément aux pluies ou à la courbe de tarage (Doctoral dissertation, Université Grenoble Alpes).

Mitchell, M. (1998). An introduction to genetic algorithms. MIT press.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. Journal of hydrology, 10(3), 282-290.

C. Perrin, C. Michel & V. Andréassian. (2007). Modèles hydrologiques du génie rural (GR).
Rapport, Edition CEMAGREF.

Sarann., Charles, C., & Degré, A. (2013). Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale. A review. Biotechnologie, Agronomie, Société et Environnement, 17(2), 392.

Schalkoff, R. J. (1997). Artificial neural networks (Vol. 1). New York: McGraw-Hill.

Shukla, S., & Lettenmaier, D. P. (2011). Seasonal hydrologic prediction in the United States: understanding the role of initial hydrologic conditions and seasonal climate forecast skill. Hydrology and Earth System Sciences, 15(11), 3529-3538.

Teuling, A. J., I. Lehner, J. W. Kirchner, and S. I. Seneviratne (2010), Catchments as simple dynamical systems : Experience from a swiss prealpine catchment, Water Resources Research, 46(10), n/a–n/a, doi :10.1029/2009WR008777.

De Vicente, J., Lanchares, J., & Hermida, R. (2003). Placement by thermodynamic simulated annealing. Physics Letters A, 317(5), 415-423.

Vrugt, J. A., C. J. F. ter Braak, M. P. Clark, J. M. Hyman, and B. A. Robinson (2008), Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with markov chain monte carlo simulation, Water Resources Research, 44(12), n/a–n/a, doi:10.1029/2007WR006720.